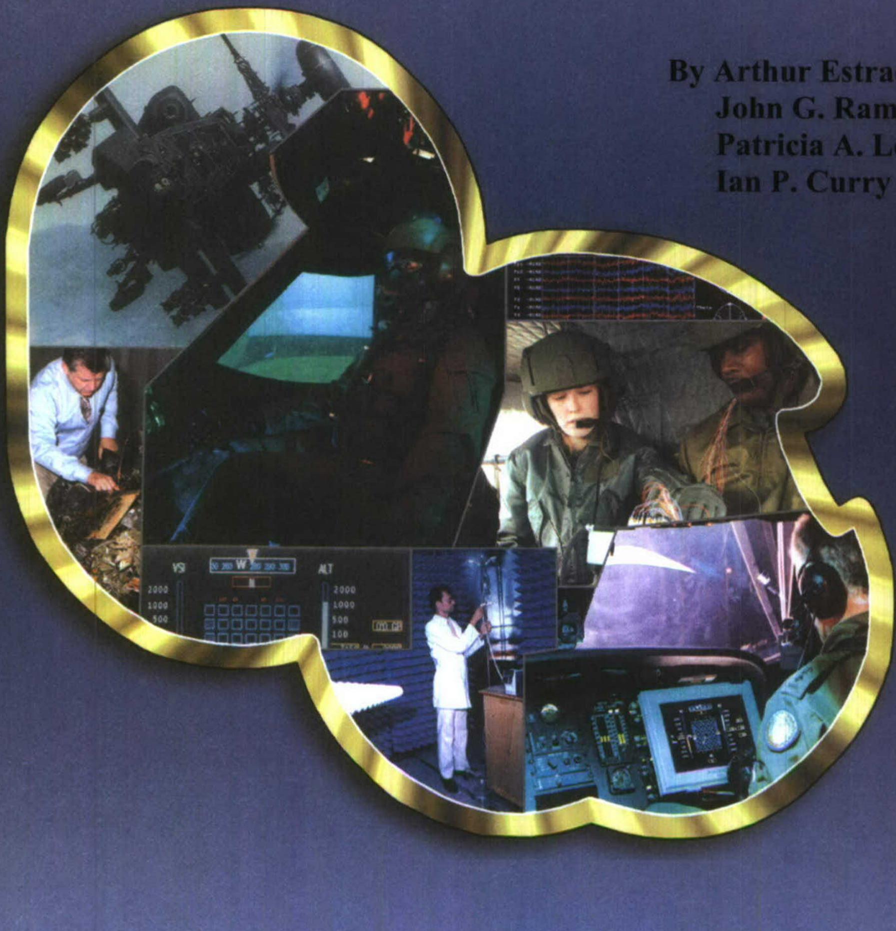
USAARL Report No. 2008-09

Performance Sustainment of Two Man Crews during 87 Hours of Extended Wakefulness with Stimulants (Dexedrine, Caffeine, Modafinil) and Napping: Analysis of Aircrew Performance during In-Flight Emergency Situations

By Arthur Estrada
John G. Ramiccio
Patricia A. Leduc
Ian P. Curry



Warfighter Performance and Health Division

May 2008

USAARL

U.S. Army
Aeromedical Research
Laboratory

Notice

Qualified requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), 8725 John J. Kingman Road, Suite 0944, Fort Belvoir, Virginia 22060-6218. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about Laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Human Use

Human subjects participated in this study after giving their free and informed voluntary consent. Investigators adhered to Army Regulation 70-25 and USAMRMC Regulation 70-25 on use of volunteers in research.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
12 May 2008

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Performance Sustainment of Two Man Crews during 87 Hours of Extended Wakefulness with Stimulants (Dexedrine, Caffeine, Modafinil) and Napping: Analysis of Aircrew Performance during In-Flight Emergency Situations

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
622787

6. AUTHOR(S)
Arthur Estrada
John G. Ramiccio
Patricia A. Leduc
Ian P. Curry

5d. PROJECT NUMBER
878

5e. TASK NUMBER
O

5f. WORK UNIT NUMBER
DA0G0165

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Aeromedical Research Laboratory
Post Office Box 620577
Fort Rucker, AL 36362-0577

8. PERFORMING ORGANIZATION REPORT NUMBER
USAARL 2008-09

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Material Command
504 Scott Street
Fort Detrick, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The objective of this effort, a small part of a larger study, was to determine the extent to which dextroamphetamine, caffeine, modafinil, and placebo affected participant ability to employ good aircrew coordination practices and function as an effective crew during emergency situations. Thirty-two UH-60 rated aviators, participating as crews of two under one of the four conditions, were assessed for aircrew coordination, crew response time, and seven flight performance measures during emergencies presented at 23, 47, and 71 hours of continuous wakefulness. While some aspects of the flight performance revealed statistically significant differences, it is clear that drug condition made no consistent significant difference in aircrew performance. The fatigue effect was limited to aircrew coordination, while not significantly affecting flight performance or response time. The variable which had the most significant albeit inconsistent effect on performance was flight experience. The findings demonstrate that analyses of performance during relatively short, stimulating experiences do not provide reliable evidence as to the effectiveness of stimulants.

15. SUBJECT TERMS
stimulants, aircrew coordination, drug effects, in-flight emergencies

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(Include area code)*
UNCLAS	UNCLAS	UNCLAS	SAR	40	Dr. Loraine Parish St. Onge / 334-255-6906

Acknowledgements

The authors would like to express their sincere gratitude to the following people for their contributions to this project.

- Mr. Larry Woodrum for developing the flight profile which supported this study.
- Ms. Elizabeth Stokes for help with all matters administrative.
- LTC Robert Wildzunas, Dr. John S. Crowley, and Dr. Loraine Parish St. Onge for their editorial assistance and steadfast support.

Table of contents

Page

Table of contents (continued)

Page

List of figures

Table of contents (continued)
List of figures (continued)

Page

List of tables

List of tables (continued)

Page

Introduction

This report presents supplemental findings of the research project *Performance Sustainment of Two Man Crews during 87 Hours of Extended Wakefulness with Stimulants (Dexedrine, Caffeine, Modafinil) and Napping* by Leduc et al. (in press). The focus of this report is on aircrew performance during in-flight emergency situations under the four treatment conditions.

Military significance

Around-the-clock military operations today are the norm rather than the exception (Miller, 2005); with night operations a significant component of combat and training missions (Comperatore et al., 1993). Military personnel often are required to work for long periods of time without rest. This lack of rest can degrade their ability to perform their duties efficiently, correctly, and in some cases, safely (Caldwell & Caldwell, 1993). Leduc et al. (in press) write that when it is not possible to obtain adequate sleep because of operational constraints, alertness-promoting compounds can be an alternative strategy. Stimulants are effective and easy to use because their utility is not dependent upon environmental manipulations or scheduling modifications. Therefore, amphetamines and other stimulants have been used extensively in military operations (Miller, 1997; Emmonson & Vanderbeek, 1993). Of the alertness-promoting compounds currently available, caffeine, dextroamphetamine (Dexedrine), and modafinil appear to hold the most promise for use in aviation operations and have been shown to be effective in a variety of situations (Akerstedt & Ficca, 1997).

Background

In the military aviation environment, where critical decisions, coordination, and task execution are frequently required under urgent adverse conditions, the effects of alertness-promoting compounds must be evaluated for their impact on aircrew coordination. Aircrew coordination is characterized by crew relationships, workload distribution, communication, and situational awareness. The importance of aircrew coordination in today's complex combat aircraft cannot be overstated. A report by the U.S. Army Combat Readiness Center (2007) states that aviation accident investigation and analysis highlighted inadequate aircrew coordination as a leading cause of aviation accidents. In fact, in the last five years, deficient aircrew coordination was cited as a contributing factor in 39% of all Army aviation accidents (Lyle, 2007).

During 2006, the Army completely overhauled its aircrew coordination training (ACT) program by instituting the Aircrew Coordination Training Enhancement (ACT-E) Program and required the use of the new Aircrew Coordination Training Aircrew Guide (Final Prototype) (United States Army Aviation Center, 2002). According to the Guide, ACT is defined as a set of principles, attitudes, procedures, and techniques that transform individuals into an effective crew. All U.S. Army graduate aviators must be qualified through the ACT-E Program in order to fly Army aircraft.

Research objective

This report reflects a small part of a larger study. As such, the objective of this aspect of the investigation was to determine the extent to which, after continuous sustained operations, the four testing conditions (dextroamphetamine, caffeine, modafinil, and placebo) allowed the participants to employ good aircrew coordination practices and function as an effective crew during emergency situations.

Methods

General

The study protocol was approved in advance by the U.S. Army Aeromedical Research Laboratory (USAARL) Human Use Committee. Each subject provided written informed consent before participating. All flights of the study were conducted in the USAARL NUH-60A simulator (Figure 1) and were videotaped for later review and analyses. As stated in Leduc et al. (in press), the simulator produces computer-generated visual displays and possesses a multi-channel, data acquisition system for analyzing various parameters of flight such as heading, airspeed, and altitude control. The flights involved a variety of maneuvers (Table 1): one hover, one VMC (Visual Meteorological Conditions) straight-and-level, one VMC DME (Distance Measuring Equipment) arc, one IMC (Instrument Meteorological Conditions) DME arc, one climb, one IMC straight and level, and an instrument landing system (ILS) approach flown under IMC. The duration of each flight was approximately 60 minutes (30 minutes per pilot on the flight controls). In other words, each participating pilot flew all of the maneuvers listed in table 1 (in that sequence) on each flight. The order in which each pilot flew first alternated with each flight.

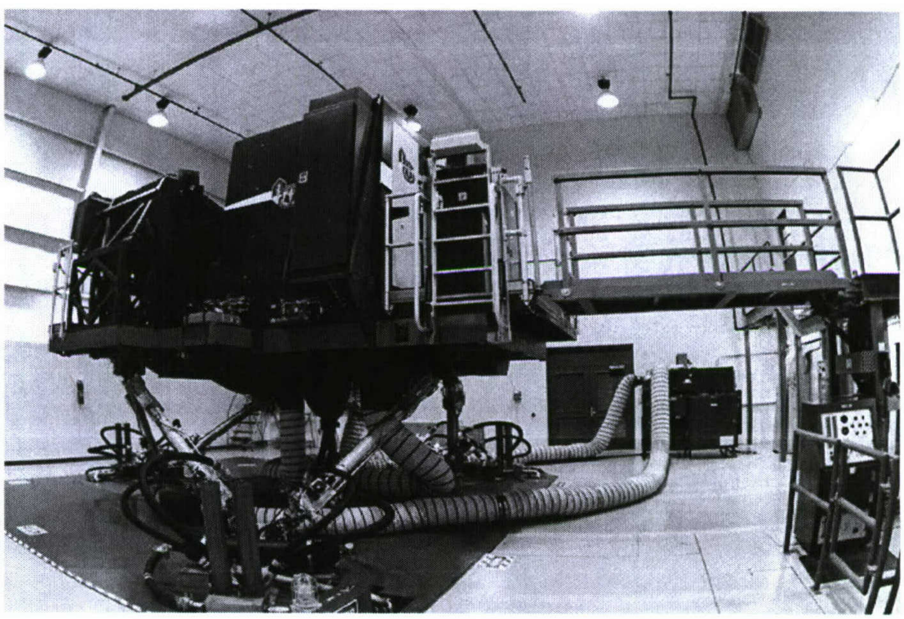


Figure 1. The USAARL NUH-60A Black Hawk helicopter simulator.

Table 1.
Flight maneuvers (Leduc et al., in press).

Maneuver	*Description*
1. Stationary hover	Perform a 10 feet (ft) stationary hover for 2 minutes (min)
2. Straight and level	Maintain VMC flight at 1500 ft, 100 knots indicated airspeed (KIAS) for 2 min
3. DME arc	Maintain 2 mile arc, 1500 ft, and 100 KIAS for 3 min
4. Climb	Climb from 1500 to 2500 ft at 500 feet per minute and 100 KIAS
5. DME arc (IMC)	Maintain 2 mile arc, 1500 ft, and 100 KIAS for 3 min
6. ILS straight and level	Maintain IMC flight at 2500 ft, 100 KIAS for 2 min
7. ILS approach	Execute ILS approach (measured from the locator outer marker to the middle marker)

The protocol's testing schedule (Appendix A) required that 22 flights be conducted over the 145-hour study period. This report describes the evaluation of only three of the 22 60-minute flights: specifically, those that occurred at 0500 hours (hr) on testing days 1, 2, and 3 (days 4, 5, and 6 of the entire protocol) during which two emergency situations were presented to the aircrews, one when each pilot was on the flight controls. The emergency situations were presented during the DME arc (IMC) maneuver (#5). In other words, it is important to note that this report represents analyses of flight and aircrew coordination performance solely during the period of time of Maneuver #5 in which the aircrews were dealing with an in-flight emergency situation.

Study population and description

As described in Leduc et al. (in press), 32 UH-60 rated aviators were recruited for this research (30 males and 2 females). Volunteers were monetarily compensated for their participation and all active duty military volunteers were on leave. The participants met all inclusion criteria and their ages ranged from 23 to 53 with 34.6 years being the population average. Their total flight experience ranged from 170 to 12,000 hr with an average of 1893 hr experience. The average participant was skilled in the UH-60 helicopter reporting 1123 flight hr. Actual UH-60 experience ranged from 55 to 5000 hr. Table 3 contains the population's distribution based on flight experience.

Procedure

For drug administration procedures, see Leduc et al. (in press). As stated earlier, all flights were divided into two 30-minute segments resulting in all of the flight maneuvers in table 1 being performed by each pilot during the course of the 1-hour flight. During the 0500 hr flight on testing days 1, 2, and 3, the simulator operator/data collector presented an emergency situation (see table 2) during Maneuver #5 of each flight segment which necessarily required the

aircrew to respond to and remedy the situation. In order to respond correctly, the aircrew had to apply the appropriate emergency procedure in an appropriate manner in accordance with the Aircrew Coordination Training Aircrew Guide (Final Prototype) (United States Army Aviation Center, 2002), the Aircrew Training Manual Utility Helicopter H-60 Series (Headquarters, Department of the Army, 2005), and the Operator's Manual for UH-60A Helicopter, UH-60L Helicopter, EH-60A Helicopter (Headquarters, Department of the Army, 2006). Note that, according to established practice, the pilot on the flight controls is responsible for aircraft control, obstacle avoidance, and the proper execution of the specific steps of an emergency procedure for the *flying* pilot. The pilot not on the controls is responsible for navigation, in-flight computations, assisting the pilot on the controls (as requested), and the proper execution of the emergency procedure steps required of the *non-flying* pilot.

Table 2.
Emergency situations.

Testing Day	*Flight Segment*	*Emergency Situation Presented*
1	1	Battery fault caution appears
	2	Stabilator malfunction – automatic mode failure
2	1	#1 generator caution appears
	2	Flight path stabilization roll trim hardover
3	1	Battery low charge caution light appears
	2	#2 generator caution appears

Evaluation of aircrew performance

As mentioned earlier, objective flight performance data were collected by the simulator's multi-channel, data acquisition system. Specialized software allowed the collection of flight parameter samples from each maneuver, computation of root mean square (RMS) error values, and computer-generated maneuver scores. Using a scoring protocol developed by Caldwell et al. (1992), these computer-generated maneuver scores were derived by first categorizing each sample of a given measure (altitude, airspeed, and DME for this study) into one of six bins ranging from worst to best scores (0, 20, 40, 60, 80, or 100 %) depending on how far that sample deviated from a predetermined standard as shown in Appendix C. At the conclusion of the first step, each bin contained one integer value which represented the number of samples classified into that particular bin. Then, the number of total samples collected on each measure during each maneuver was determined. The number of samples in each bin was multiplied by the weighing factor for the respective bin (0, 20, 40, 60, 80, or 100%); the results were then summed

and divided by the number of measures (in this case three, for altitude, airspeed, and DME). Thus, at the completion of this entire procedure, there was one performance score per maneuver.

Two highly-experienced, ACT-E-qualified, research instructor pilots, blinded to the experimental conditions, viewed and listened to the videotaped recordings of the flights in which emergency situations occurred and they assessed the aircrews' coordination performance. The flights were recorded in split screen presenting frontal views of each pilot, a view of the central cockpit, and a view of the outside scene (figure 2). The assessments were made based on the Aircrew Coordination Training Aircrew Guide's (2002) grading guidance and by using the Aircrew Coordination Training Performance Evaluation Checklist (Appendix B). Upon review of the emergency response, the assessing pilots determined to what extent the flight crews met the ACT-E's Crew Coordination Objectives (CCOs) and Basic Qualities (BQs) (see Appendix B for a listing of CCOs and BQs). The grading guidance provided direction in determining performance on a scale of 1 to 7 (table 3). For the purposes of this report, the assessed performance grades will be referred to as *aircrew coordination scores*.

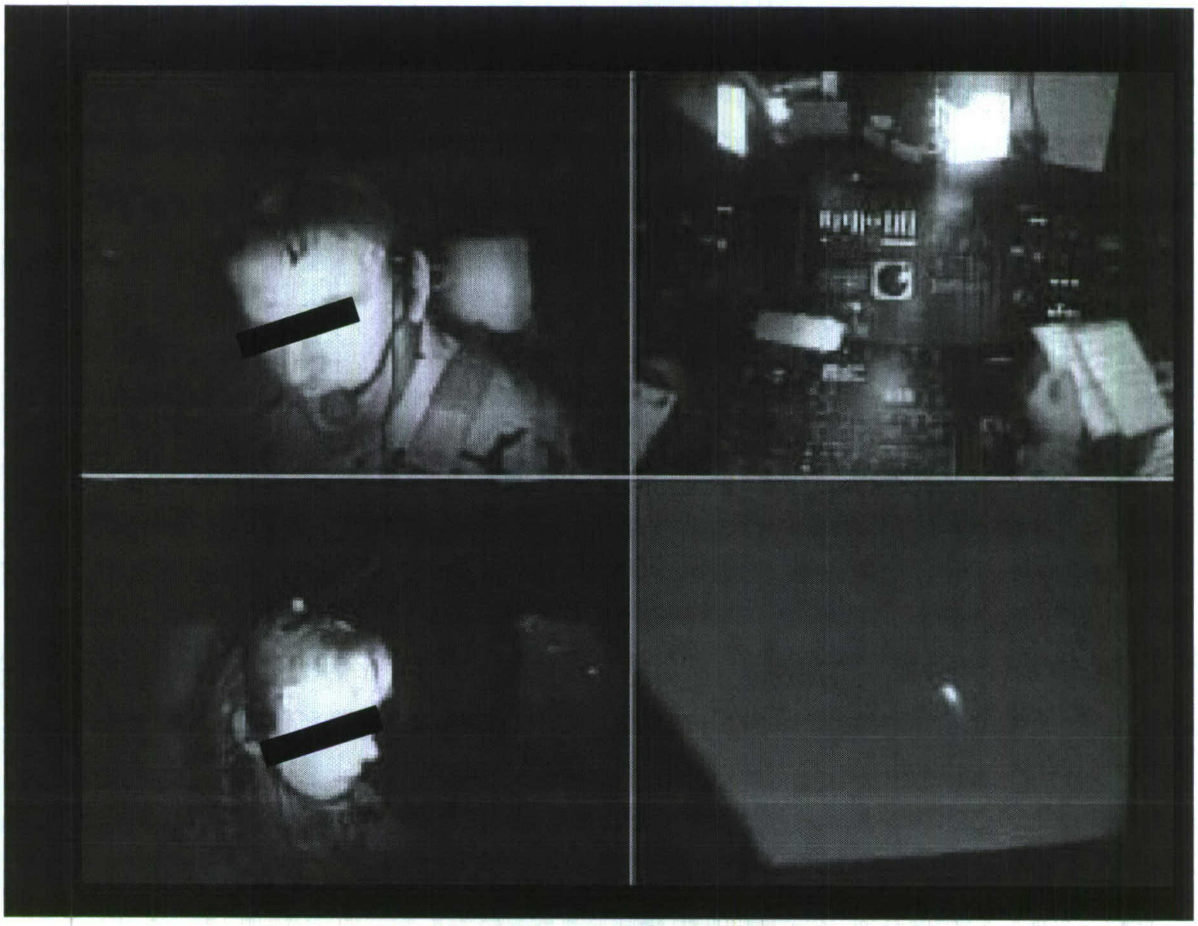


Figure 2. Sample of flight videotaped recordings.

Table 3.
Aircrew coordination scores.

1 = Below standards
2 = Frequently below standards
3 = Occasionally below standards
4 = Meets standards
5 = Occasionally above standards
6 = Frequently above standards
7 = Above standards

Results

All statistical analyses were conducted using SPSS® 12.0 with statistical significance set at an alpha level of .05 for all statistical tests. Independent variables consisted of drug condition, flight day, and aircrew and individual flight experience. The dependent variables were aircrew coordination scores, individual flight performance measure scores (for maintaining altitude, airspeed, and the DME arc), the maneuver score (an aggregate of the three individual measures), the root mean square error (RMSE) for each measure, and the latency to detect and correct the situation. Flight experience was categorized (table 4) into groups in order to facilitate data analysis.

Table 4.
Flight experience categories and population distributions.

Group	Total Flight Time (hr)	Number of Participants	Total UH60 Flight Time (hr)	Number of Participants
1	1 – 1000	8	1 – 500	9
2	1001 – 2000	18	501 – 1000	5
3	2001 – 3000	2	1001 – 1500	12
4	3001 – 4000	2	1501 – 2000	4
5	> 4000	2	> 2000	2

An initial exploration of the data found that all but the aircrew coordination scores (with a skewness statistic of -.046) of the collected data sets were significantly skewed (not normally distributed) and thus, did not meet the basic assumptions of parametric tests. Efforts to normalize the distribution of the data through transformations were unsuccessful. As such, appropriate nonparametric tests were employed to analyze those data. Parametric tests were used to analyze aircrew coordination.

Aircrew coordination scores

During this study, 27 of the 96 flights could not be assessed for aircrew coordination during the emergency situation due to a failure of either the recording equipment or of the simulator operator to present the emergency situation. As mentioned above, the distribution of the overall aircrew coordination performance scores (ratings) was normal as evidenced by table 5.

Table 5.
Aircrew coordination scores descriptive statistics.

	Frequency	Percent
Below standards	7	7.3
Frequently below standards	7	7.3
Occasionally below standards	20	20.8
Meets standards	22	22.9
Occasionally exceeds standards	7	7.3
Frequently exceeds standards	6	6.3
Total	69	71.9
Missing	27	28.1
Total	96	100.0

In order to gain meaningful insight into what aspects of aircrew coordination enhanced or hindered performance, an examination of the frequency at which the aircrew coordination basic qualities contributed to the flight aircrew coordination scores was useful. A review of table 6 shows that aircrew that performed above the established aircrew coordination standard communicated clearly in a timely manner, maintained situational awareness, and acknowledged actions. Those who performed below the standards not only performed those tasks poorly, but were especially noted for mismanagement of unexpected events (i.e., emergency situations), for inefficiently prioritizing actions and distributing workload, and for inadequately cross-monitoring the actions of the other pilot.

Table 6.

Frequency of Basic Qualities cited as contributing to aircrew coordination scores.

	Basic Qualities (see legend below)												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1 – Below standards	1			1	2	5	4	4	1	5			
2 – Frequently below standards				1	2	2	1	4	1	2	1	1	
3 – Occasionally below standards	2			3	6	7	1	10			1		
4 – Meets standards													
5 – Occasionally above standards				2		5	3	4		2			
6 – Frequently above standards			1			5	4	4	1	2	1		
7 – Above standards													

Legend- Basic Qualities

1 Establish and maintain team leadership and crew climate
2 Premission planning and rehearsal accomplished
3 Application of appropriate decision making techniques
4 Prioritize actions and distribute workload
5 Management of unexpected events
6 Statements and directives clear, timely, relevant, complete and verified
7 Maintenance of situational awareness
8 Decisions and actions communicated and acknowledged
9 Supporting information and actions sought from crew
10 Crewmembers actions mutually cross-monitored
11 Supporting information and actions offered by crew
12 Advocacy and assertion practiced
13 Crew/flight after-action reviews accomplished

Drug effects on performance during emergency situations

Aircrew coordination

In order to test for differences in aircrew coordination by drug condition, a Kruskal-Wallis test (nonparametric alternative to an ANOVA) was conducted. The results indicated that aircrew coordination did not significantly differ with drug condition [$\chi^2(5, N = 69) = 9.722, p < .084$]]. Figure 3 provides a look at the distribution by the frequency of aircrew coordination scores per drug condition. Unexpectedly, it was the placebo group which achieved the most above standard scores (eight times) compared to caffeine (once), dextroamphetamine (thrice), and modafinil (once) (table 7, shaded columns) .

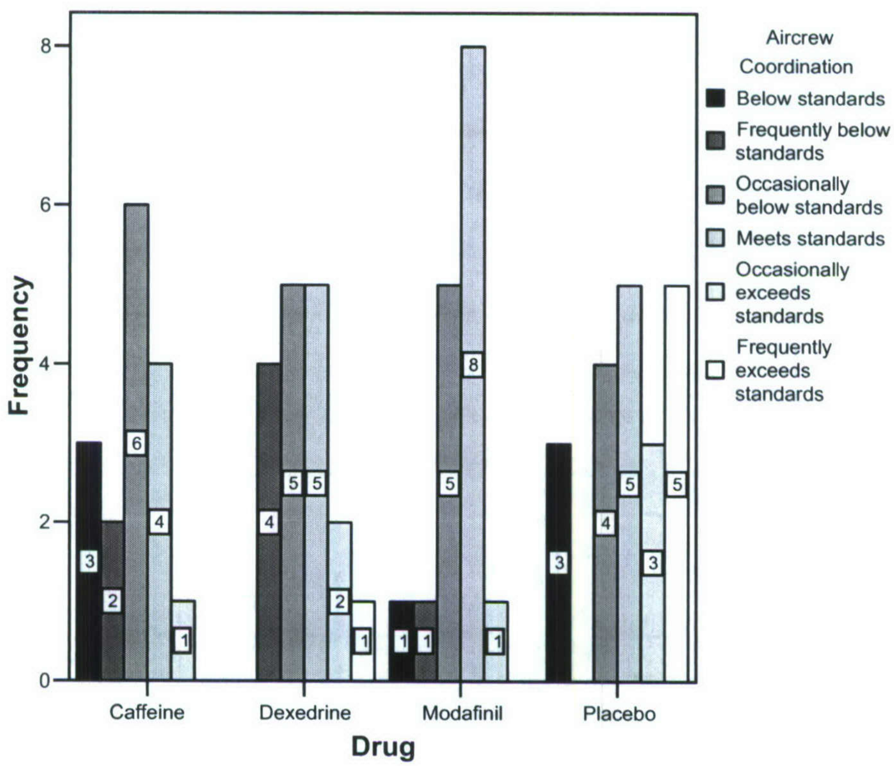


Figure 3. Frequency of aircrew coordination scores by drug condition.

Flight performance

The Kruskal-Wallis test was again used to test for the effects of drug condition on flight performance during the emergency situations. The results of these tests (tables 7, 8, and 9) indicate that the only significant group difference between drug conditions existed in the DME arc RMSE. The modafinil condition differed from the other groups with a significantly large majority (19 of 24) of having RMSEs less than the group median (table 8, shaded). Clearly stated, the modafinil pilots maintained a more precise DME arc than those of the other drug

conditions during the distractive emergency situation. For a basic comparison, figure 4 presents the DME arc RMSE means by drug condition. Note the minimal average variation from zero (0.07) for the modafinil pilots. Although significant, this isolated finding is not sufficient to declare modafinil superior in maintaining flight performance. A table containing the descriptive statistics of all performance scores by drug condition is available in Appendix D.

Table 7.
Kruskal-Wallis test for drug condition and performance.

	Altitude Score	Airspeed Score	DME arc Score	Total Maneuver Score	Altitude RMSE	Airspeed RMSE	DME arc RMSE
Chi-Square	3.233	2.896	7.015	5.235	2.472	2.466	10.026
df	3	3	3	3	3	3	3
Significance	.357	.408	.071	.155	.480	.481	.018*

Grouping variable: Drug condition; * p < .05; ** p < .01.

Table 8.
Median test for drug condition and performance.

		Drug			
		Caffeine	Dextroamphetamine	Modafinil	Placebo
Altitude Score	> Median	8	14	15	10
	<= Median	16	10	9	12
Airspeed Score	> Median	7	13	13	14
	<= Median	17	11	11	8
DME Arc Score	> Median	8	11	16	11
	<= Median	15	13	8	11
Total Maneuver Score	> Median	8	12	14	12
	<= Median	15	12	10	10
Altitude RMSE	> Median	14	10	11	12
	<= Median	10	14	13	10
Airspeed RMSE	> Median	14	13	11	9
	<= Median	10	11	13	13
DME Arc RMSE	> Median	15	12	5	12
	<= Median	8	12	19	10

Table 9.
Median test statistics for drug condition and performance.

	Altitude Score	Airspeed Score	DME Arc Score	Total Maneuver Score	Altitude RMSE	Airspeed RMSE	DME Arc RMSE
N	94	94	93	93	94	94	93
Median	70.13	86.59	87.11	79.72	25.10	1.95	.08
Chi-Square	5.015	6.136	4.954	2.969	1.682	1.727	10.240
df	3	3	3	3	3	3	3
Significance	.171	.105	.175	.397	.641	.631	.017*

Grouping variable: Drug condition; * p < .05; ** p < .01.

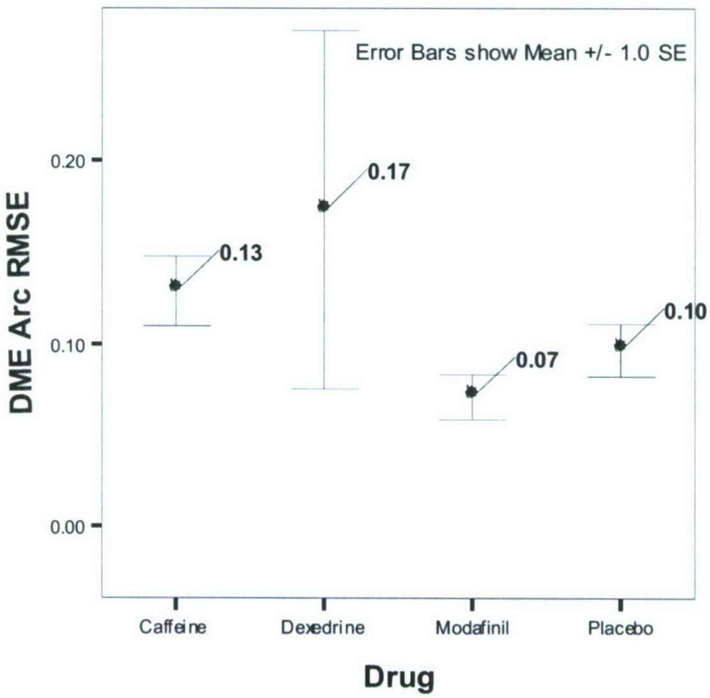


Figure 4. DME arc RMSE mean comparisons.

Response to the emergency situation

The initial analysis, which included an extreme outlier of 9 seconds in the dextroamphetamine group, showed that the drug condition made no significant difference in emergency response times [F (3, 68) = .377, p = .770] (figure 5). However, following the Winsorization of the outlier, the dextroamphetamine group demonstrated significantly faster response times when compared to the other condition groups [F (3, 68) = 3.313, p = .025] (figure 6).

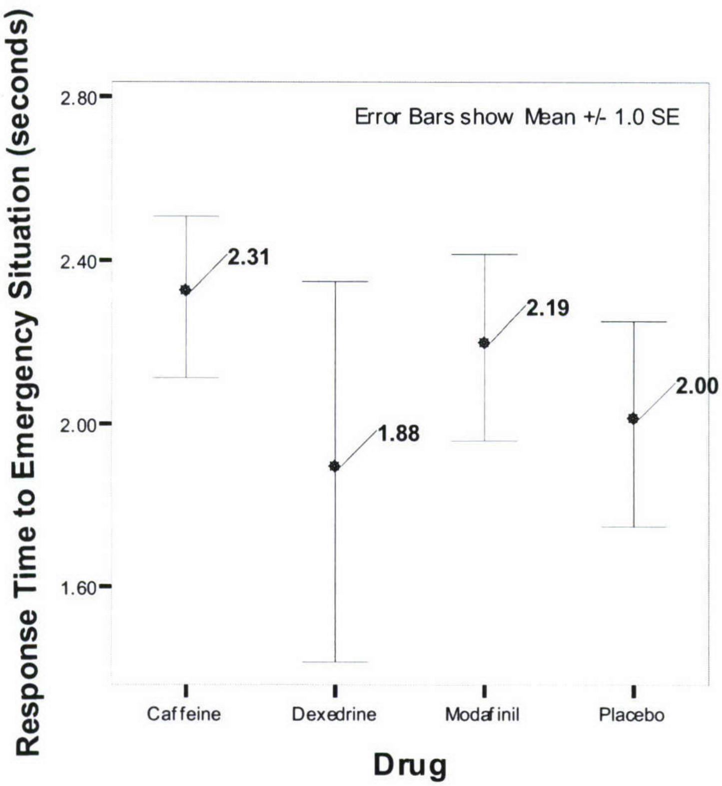


Figure 5. Emergency response times by drug group (outlier included)

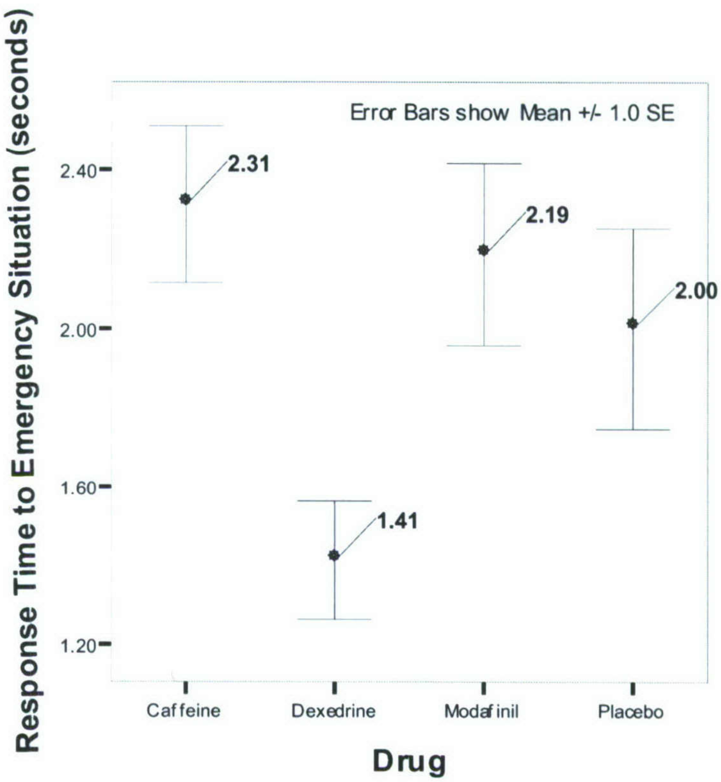


Figure 6. Emergency response times by drug group (outlier Winsorized).

Fatigue effects on performance during emergency situations

As stated earlier, the analyzed flights occurred on days 4, 5, and 6 of the study. As a result of the testing schedule, participants had experienced 23 hr of continuous wakefulness by the 0500 flight on day 4, 47 hr by the 0500 flight on day 5, and 71 hr by the 0500 flight on day 6. It was reasonable to expect differences in human performance from day 4 to day 6 due to the effects of fatigue.

Aircrew coordination

Analysis by ANOVA of the aircrew coordination during the flight emergency situations showed a significant decline aircrew coordination performance between the flight days [F (2, 68) = 3.900, p = .025]. Figure 7 illustrates the deterioration in average aircrew coordination performance as wakefulness continued. Figure 8 presents the distribution of the frequency of aircrew coordination score categories (from below standards to frequently exceeds standards) per flight day. Note the increase in the number (frequency) of below standard assessments as the days passed.

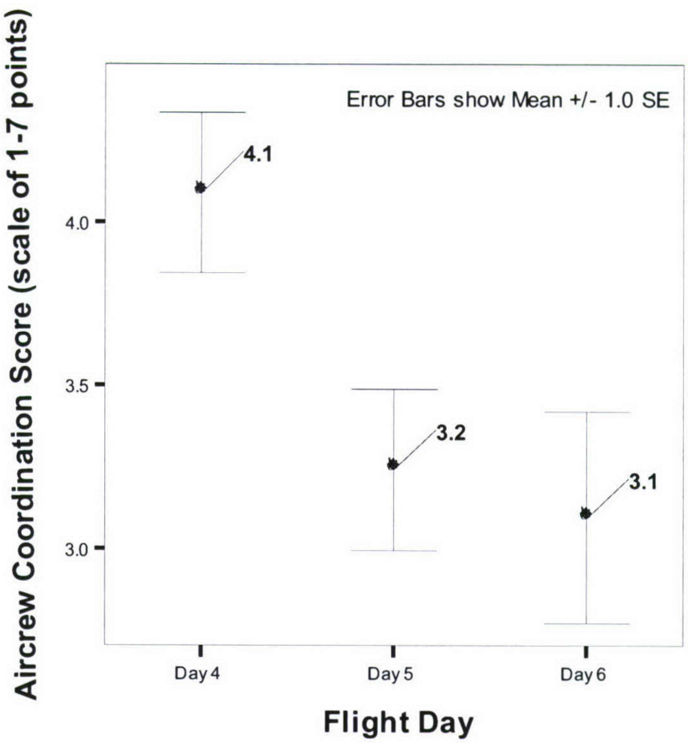


Figure 7. Aircrew coordination scores by flight day.

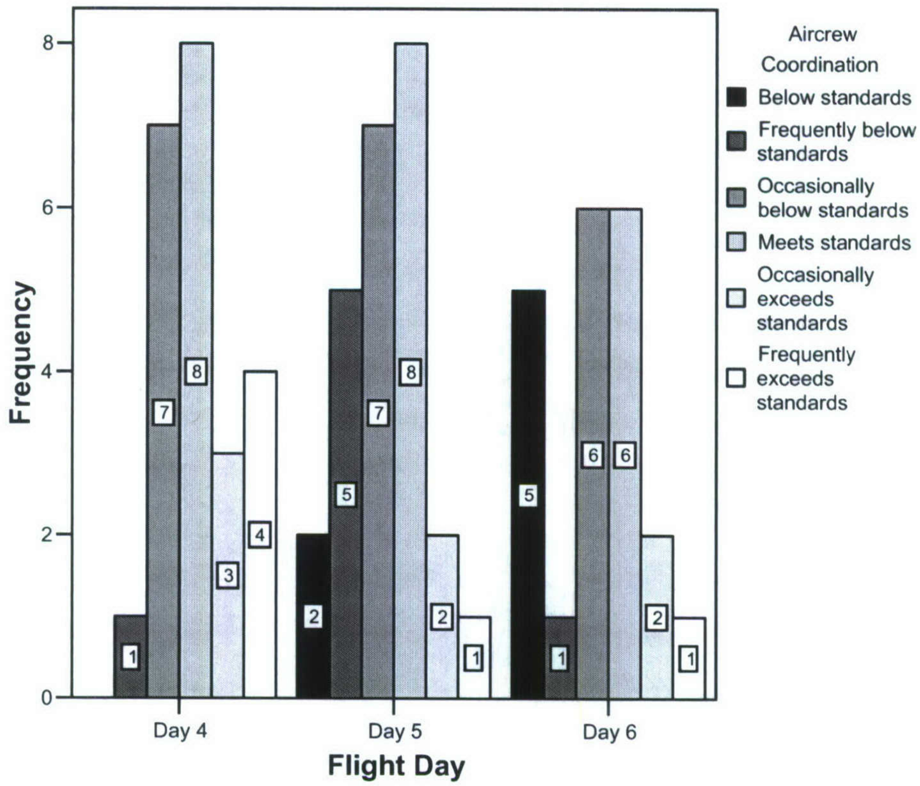


Figure 8. Frequency of aircrew coordination scores by flight day.

Flight performance

In contrast, flight performance during the emergency situations did not significantly differ over time. Analyses by Kruskal-Wallis tests resulted in no significant differences discovered in the flight performance data over the three days (table 10).

Table 10.
Kruskal-Wallis test for test day and performance.

	Altitude Score	Airspeed Score	DME arc Score	Total Maneuver Score	Altitude RMSE	Airspeed RMSE	DME arc RMSE
Chi-Square	2.285	3.994	.152	2.037	2.938	1.717	.267
df	2	2	2	2	2	2	2
Significance	.319	.136	.927	.361	.230	.424	.875

Grouping variable: Test day; * $p < .05$; ** $p < .01$.

Response to the emergency situation

An analysis of variance showed that flight day made no significant difference in emergency response times [$F(2, 68) = .514, p = .600$] when the data included an aircrew in the dextroamphetamine drug group which took 9 s (extreme outlier) to respond to a battery low charge caution light on day 6 (figure 9). Attempts to Winsorize the outlier made no difference in significance.

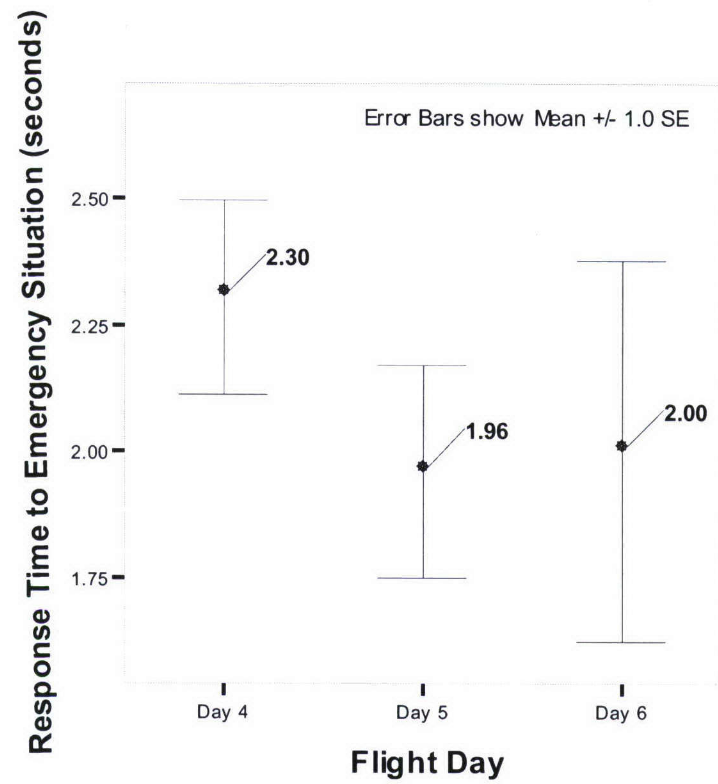


Figure 9. Emergency response times by flight day (outlier included).

Flight experience effects on performance during emergency situations

In the following analyses, experience is characterized as total flight experience and as flight experience in the UH-60 helicopter. Experience in specific helicopter models implies familiarity with the aircraft's handling qualities and cockpit environment which can be a factor in an aviator's flight performance, especially during critical and unexpected situations.

Total flight experience

Aircrew coordination

An ANOVA was conducted to determine possible differences in aircrew coordination scores based on the total combined aircrew flight experience. No significant differences were noted [F (3, 68) = .031, p = .993]. In contrast, the total flight experience of the pilot actually on the flight controls did make a significant difference in aircrew coordination scores [F (4, 68) = 2.913, p = .028]. A post hoc test (Scheffe's) highlighted the significant difference (p = .03) between the quality of the aircrew coordination when the flying pilots had 2001-3000 hr versus those with 3001-4000 hr. When the flying pilot had 3001-4000 hr, the aircrew coordination scores were the lowest average of all experience groups (figure 10). A subsequent test determined that it did not matter if the pilot flying was the higher- or lower-time pilot of the aircrew [F (1, 68) = .064, p = .801].

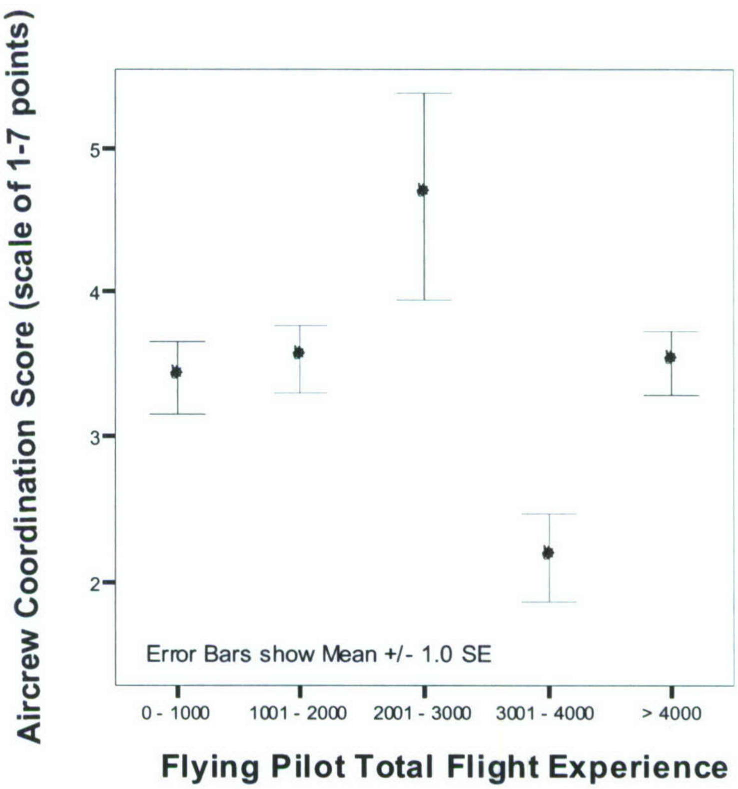


Figure 10. Aircrew coordination scores by total flight experience (hr) of pilot flying.

Flight performance

Upon subjecting all of the performance measures to the Kruskal-Wallis test, table 11 shows that when total (combined) crew flight experience was considered, there were significant group differences in altitude, airspeed, total maneuver, and altitude RMSE. The subsequent Median test (table 12, shaded) and statistics (table 13) show that, as a group, the crews with greater than 4000 hr of combined flight experience actually performed significantly more poorly than the other groups on the four measures. This was evidenced by the larger percentage of below-median scores for altitude, airspeed, and the total maneuver, and by the above-median altitude RMSE compared to the other experience groups.

Table 11.
Kruskal-Wallis test for total fight experience and performance.

	Altitude Score	Airspeed Score	DME arc Score	Total Maneuver Score	Altitude RMSE	Airspeed RMSE	DME arc RMSE
Chi-Square	13.057	8.271	3.432	11.591	14.049	5.909	3.651
df	3	3	3	3	3	3	3
Significance	.005**	.041*	.330	.009**	.003**	.116	.302

Grouping variable: Total crew flight experience; * $p < .05$; ** $p < .01$.

Table 12.
Median test for total crew fight experience and performance.

		Total Flight Experience				
		0 - 1000	1001 - 2000	2001 - 3000	3001 - 4000	> 4000
Altitude Score	> Median	0	12	12	18	5
	<= Median	0	18	11	5	13
Airspeed Score	> Median	0	11	14	16	6
	<= Median	0	19	9	7	12
DME arc Score	> Median	0	15	14	9	8
	<= Median	0	14	9	14	10
Total Maneuver Score	> Median	0	12	15	15	4
	<= Median	0	17	8	8	14
Altitude RMSE	> Median	0	16	10	5	16
	<= Median	0	14	13	18	2
Airspeed RMSE	> Median	0	17	10	8	12
	<= Median	0	13	13	15	6
DME arc RMSE	> Median	0	11	10	13	10
	<= Median	0	18	13	10	8

Table 13.
Median test statistics for total crew fight experience and performance.

	Altitude Score	Airspeed Score	DME arc Score	Total Maneuver Score	Altitude RMSE	Airspeed RMSE	DME arc RMSE
N	94	94	93	93	94	94	93
Median	70.13	86.59	87.110	79.717	25.10	1.95	.0800
Chi-Square	12.147	8.742	2.420	10.669	18.761	5.055	2.433
df	3	3	3	3	3	3	3
Significance	.007**	.033*	.490	.014*	.000**	.168	.488

Grouping variable: Total crew flight experience; * $p < .05$; ** $p < .01$.

A follow-on Kruskal-Wallis test was conducted to determine whether flight performance varied depending on which pilot, the least or more experienced, of the aircrew was on the flight controls. The results indicated no significant differences in performance scores (Table 14).

Table 14.
Kruskal-Wallis test for fight performance: Least or most experienced aircrew flying.

	Altitude Score	Airspeed Score	DME arc Score	Total Maneuver Score	Altitude RMSE	Airspeed RMSE	DME arc RMSE
Chi-Square	2.105	1.201	.015	.065	1.281	.115	.191
df	2	2	2	2	2	2	2
Significance	.349	.549	.993	.968	.527	.944	.909

Total UH-60 flight experience

Aircrew coordination

Unlike total flight experience, the aircrew's combined UH-60 experience had a statistically significant effect on aircrew coordination [F (3, 68) = 3.257, p = .027], as did the flying pilot's total UH-60 experience [F (4, 68) = 2.591, p = .045]. Generally speaking, aircrew coordination was best when the most experienced UH-60 aviators were on the controls of the aircraft with the converse being equally apparent (figures 11 and 12). As in the case of total overall flight experience, it did not matter whether it was the least or most experienced UH-60 pilot of the aircrew flying the aircraft [F (1, 68) = .156, p = .694]; only that they were UH-60 experienced.

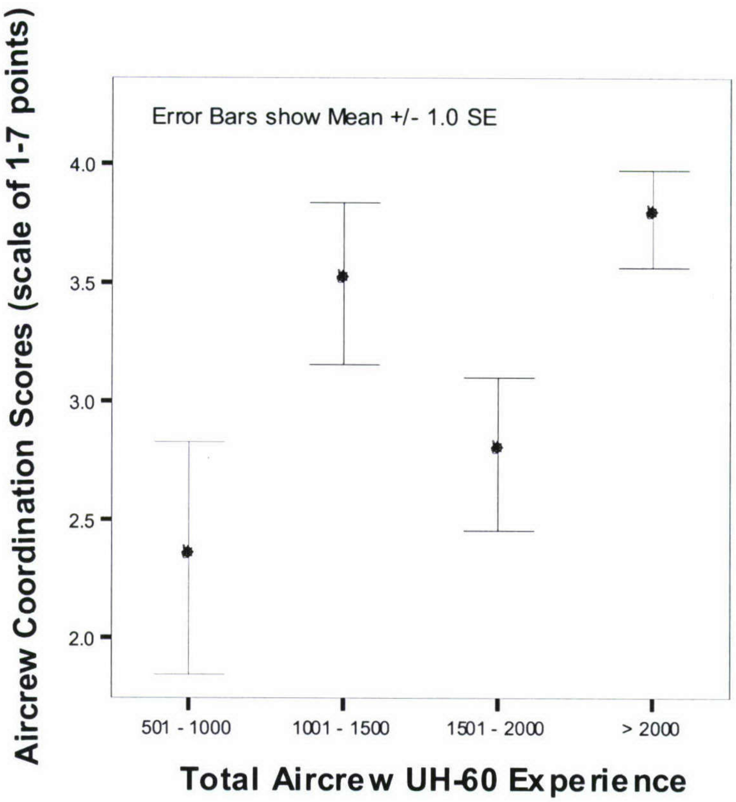


Figure 11. Aircrew coordination scores by total aircrew UH-60 flight experience.

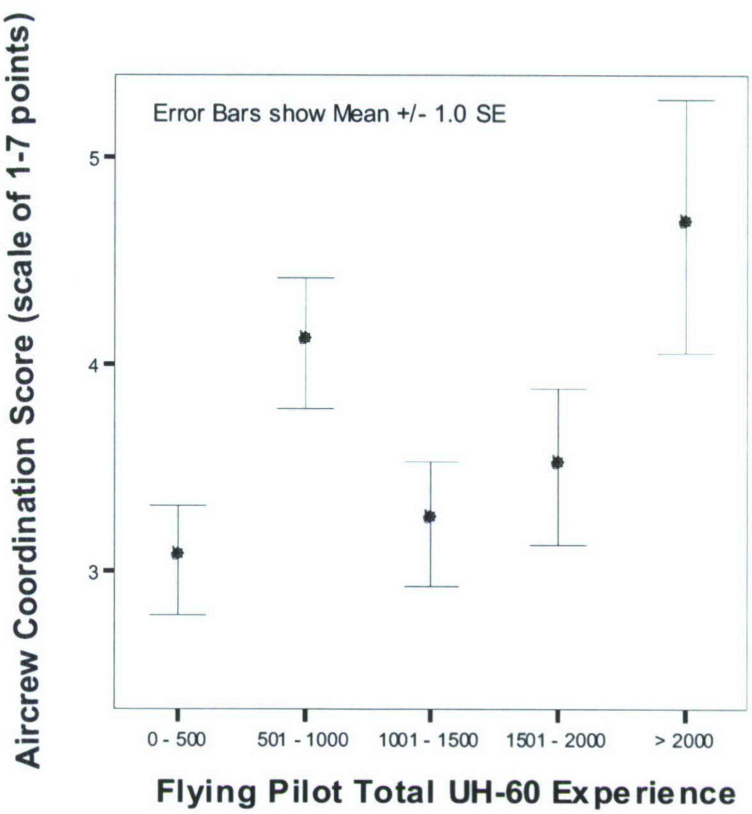


Figure 12. Aircrew coordination scores by total UH-60 flight experience of pilot flying.

Flight performance

Following the same Kruskal-Wallis procedures described previously, total crew UH-60 flight experience showed a significant difference in only one flight performance measure: that of the DME arc RMSE [H (3) = 9.120, p = .028]. The Median test indicated that combined aircrew UH-60 experience in the 1001-1500 hr range produced a high percentage of scores below the group median (16 of 23), indicating low performance errors by the group majority. In contrast, the majority of those in the other experience groups had higher DME arc RMSEs than the group median.

Also, UH-60 flight experience by the pilot on the flight controls did not impact any of the performance measures except the airspeed score [H (4) = 9.966, p = .041]. The Median test and statistics confirmed that aviators having 1001-1500 and 1501-2000 hr maintained their airspeeds more precisely (majority above overall group median) than aviators in the other experience groups [χ^2 (4, N = 94) = 9.542, p = .049]. Correspondingly, an additional Kruskal-Wallis test showed a significant difference in airspeed scores when the more experienced pilot of the flight crew was on the controls [H (1) = 6.937, p = .008]. The Median test statistics showed that the more experienced of the aircrew maintained airspeed significantly more precisely than the less experienced during the emergency situations [χ^2 (1, N = 94) = 4.257, p = .039].

Discussion

This report presents the findings of a limited facet of a larger study which examined the effects of 87 hr of extended wakefulness on the performance of two man crews taking three different stimulants (dextroamphetamine, caffeine, and modafinil) and a placebo. The specific focus of this effort was to evaluate the effects of the experimental conditions on aircrew coordination and flight performance during simulated emergency situations over three days (days 4, 5, and 6) of the seven day study. The emergency situations always occurred during the 0500-0600 flight period.

Fatigue effects

A decline in aircrew performance was expected due to the effects of fatigue. The influence of sleep loss and fatigue on neurobehavioral performance is well documented (Lamond & Dawson, 1999; Gillberg et al., 1994). Lamond and Dawson found that approximately 20-25 hr of wakefulness produced performance decrements equivalent to those observed at a blood alcohol concentration of 0.10%. Therefore, it is reasonable to expect that the periods of wakefulness under this examination (23, 47, and 71 hr) would continue to degrade performance equivalent to even higher levels of alcohol intoxication.

In this study, the data failed to show a significant effect of fatigue on flight performance or reaction times during the relatively short emergency situations. The lack of significant flight performance decrements over the three day period was surprising. This finding is consistent with Leduc et al.'s findings in the overall study (in press). In their words, "it is possible that the new element of crew interaction during the simulated flight profiles [and emergency situations] may have injected an element of stimulation and excitement which was missing in previous single-subject/pilot research flights (Caldwell et al., 1994) during which significant differences in flight performance were noted." In addition, we believe the handling characteristics of the UH-60 Black Hawk (and its simulator) may have contributed to the lack of a significant fatigue effect. The flying controls are augmented with stabilizing aircraft systems, specifically, the automatic flight control system (AFCS) and the flight path stabilization system (FPS), which can be relied upon by the flying pilot to hold flight parameters (e.g., airspeed, heading, and attitude). During the short emergency situations, the flying pilots may have relied on such stabilizing systems to maintain the flight parameters within boundaries which failed to vary enough to show significant differences over time. Similarly, the range in reaction times over the three flights was so small as to prove insignificant.

In contrast, the expected fatigue-related performance declines were clearly evident in the performance of aircrew coordination which showed statistically significantly degradation, especially from day 4 to day 5. When fatigued, flight experience (crew and individual) in the UH-60 impacted the quality of communication and coordination more so than the combined crew flight time. This may be due to pilot familiarity with the UH-60's cockpit environment, its standardized operational aspects, and the heightened emphasis on aircrew coordination in UH-60 training programs and manuals (Headquarters, Department of the Army, 2005).

Drug condition

There were no significant differences in aircrew coordination scores by drug condition, although the placebo group had a nonsignificant trend toward higher scores.

As for flying performance, even though the "modafinil pilots" maintained a significantly more precise DME arc than those under the other drug conditions, they demonstrated no superiority in any of the other flight tasks. When response to the emergency situation was assessed, the dextroamphetamine group's average response was significantly faster. The lack of consistent drug effect may be due to the cognitively arousing nature of the emergency situation. It may be the case that examination of a less stimulating portion of the flight (e.g., straight and level segments) would yield group differences based on drug condition.

Flight experience

Response times to the emergency situations did not vary significantly with flight experience. However, there was a significant interaction between total flight experience of the pilot on the controls and aircrew coordination performance. The pilots on the controls having 2001-3000 hr of experience performed significantly better than the pilots of other experience groups and the flying pilots possessing 3001-4000 hr of experience performed aircrew coordination much more poorly. Although initially interesting, a closer examination of the population data showed that these two groups were comprised of just two individuals each, hardly a representative sample of the general population. Other experience groups were more suitably represented (table 3). This makes the meaningfulness of this finding quite questionable and therefore, no generalizable inferences are made.

As for flight performance, it was discovered that aircrews with a total combined crew experience of greater than 4000 flight hr were significantly inferior in maintaining the established flight standards for altitude and airspeed, and thus, attained lower total maneuver scores and higher altitude RMSE values when compared to the other experience levels. Although inferior, their performance never placed the aircraft in an unsafe flight condition. Discussions with two senior instructor pilots, Mr. Woodrum and Mr. Ramicco (personal communication, 9 August 2007), were revealing. In general, aircrews having more experience are more confident and relaxed when dealing with adverse events. Conversely, less experienced aircrews, particularly those with moderate combined experience (i.e., 1000 to 2000 hr), tend to be more deliberate about proving their ability to handle adverse situations while simultaneously maintaining precision flight. The general lax, untroubled attitude of the more experienced aircrews, especially when dealing with the relatively benign, non-critical emergency situations presented in this study, is suspected to have contributed to their casual approach in maintaining the aircraft to exacting standards and their poor flight performance when compared to the other experience groups.

Study limitations

Despite empirical and anecdotal evidence of fatigue effects on operational performance in the real world, the absence of such effects in this study may be a consequence of the research environment. Even though simulators continue to be used during training, research, and pilot certification, the predictive relationship of data obtained from simulators and its relevance to the operational reality of flight remains inconclusive (Schmeisser, Fuller, & van de Pol, 2008). Although assessing pilot performance in the simulator provides valuable information and cost savings, performance in the simulator should not be taken as the sole predictor of performance in the aircraft.

Conclusions

The objective of this project was to determine the extent to which the four testing conditions (dextroamphetamine, caffeine, modafinil, and placebo) affected the participants' ability to employ good aircrew coordination practices and function as an effective crew during emergency situations. While some aspects of the flight performance revealed statistically significant differences, it is clear that drug condition made no consistent significant difference in aircrew performance during the emergency situations. The effect of fatigue was limited to aircrew coordination, while not significantly affecting flight performance or response time. The variable which had the most significant albeit inconsistent effect on performance was flight experience.

Also, the findings demonstrate that analyses of performance during relatively short, stimulating experiences, such as the five-minute emergency situations in this study, may not be a reliable way to assess the effectiveness of stimulants. A clearer picture of the drug effects on aircrew coordination may emerge by expanding the analysis of aircrew coordination across the entire one-hour flight.

References

Akerstedt, T., and Ficca, G. 1997. Alertness-enhancing drugs as a countermeasure to fatigue in irregular work hours. Chronobiology International. 14(2): 145-158.

Caldwell, J.L. and Caldwell, J.A. 1993. A comparison of sleep scored from electroencephalography to sleep scored by wrist actigraphy. Fort Rucker, AL: United States Army Aeromedical Research Laboratory. Report No. 93-32: 3.

Caldwell, J.L., Jones, H.D., Carter, D.J., and Caldwell, J.A. 1992. The relationship between computer scoring and safety-pilot grading of flight performance. Fort Rucker, AL: United States Army Aeromedical Research Laboratory. Report No. 92-30.

Caldwell, J. A., Caldwell, J. L., Crowley, J. S., Darling, S. R., Wallace, S., Woodrum, L. C., and Jones, H. D. 1994. Sustaining helicopter pilot performance with Dexedrine: A UH-60 simulator study. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 94-93.

Comperatore, C.A., Chiaramonte, J.A., Pearson, J.Y., Stone, L.W., Hess, G.L., and Boley, K.O. 1993. Activity/rest patterns of instructor and rated student pilots during rapid transitions from daytime to nighttime duty hours at Eastern Army Aviation Training Site. Fort Rucker, AL: United States Army Aeromedical Research Laboratory. Report No. 93-16: 7.

Emmonson, D. L. and Vanderbeek, R. D. 1993. The use of Dextroamphetamine in support of tactical air operations during Operation Desert Shield/Storm. Aviation, Space, and Environmental Medicine. 64(5): 421.

Gillberg, M., Kecklund, G., and Akerstedt, T. 1994. Relations between performance and subjective ratings of sleepiness during a night awake. Sleep. 17(3): 236-241.

Headquarters, Department of the Army. 2005. Aircrew Training Manual Utility Helicopter H-60 Series (Training Circular 1-237). Washington, DC: U.S. Government Printing Office.

Headquarters, Department of the Army. 2006. Operator's Manual for UH-60A Helicopter, UH-60L Helicopter, EH-60A Helicopter. (Technical Manual 1-1520-237-10). Washington, DC: U.S. Government Printing Office.

Lamond, N. and Dawson, D. 1999. Quantifying the performance impairment associated with fatigue. Journal of Sleep Research. 8: 255-262.

Leduc, P.A., Rowe, T.L., Martin, C.R., Curry, I.P., Wildzunas, R.M., Darlington, K.H., Dueringer, L.L., Bass, J.M., Palacio, L.V., Rouse, T., Milam, L., Childress, J., West, A., Hunt, M., Gonzales, E., McKeon, J. and Dowler, T. In press. Performance sustainment of two man crews during 87 hours of extended wakefulness with stimulants (Dexedrine, caffeine, Modafinil) and napping. Fort Rucker, AL: United States Army AeromedicalResearch Laboratory.

Lyle, C.D. 27 April 2007. In an email response to a question concerning the percentage of aviation accidents in the past five years in which deficient aircrew coordination was a contributing factor. Engineering Research Psychologist, U.S. Army Combat Readiness Center, Fort Rucker, AL.

Miller, D.B. 2005. Sleep and predicted cognitive performance of new cadets during cadet basic training at the United States Military Academy. Monterey, CA: Naval Postgraduate School.

Miller, M.A. 1997. History and epidemiology of amphetamine abuse in the United States. In Klee H (Ed.), Amphetamine misuse: International perspectives on current trends, 113-134. The Netherlands: Harwood Academic Publishers.

Schmeisser, E.T., Fuller, D.R., and van de Pol. 2008. Predictability of pilot performance from simulated to real flight in the UH-60 (Black Hawk) Helicopter. Fort Rucker, AL: United States Army Aeromedical Research Laboratory. Report No. 2008-06.

United States Army Aviation Center. 2002. Aircrew coordination training aircrew guide (final prototype). Fort Rucker, AL: Dynamics Research Corporation.

United States Army Combat Readiness Center. n.d. Fiscal Year 2006 Army safety performance. Available April 12, 2007, at https://crc.army.mil/Report/Fy06yearend.doc.

Woodrum, L.C. and Ramiccio, J.G. 9 Aug 2007. Personal interviews concerning the inferior flight performance of aircrews with greater than 4000 hours of total combined flight experience.

Appendix A

Testing schedule (Leduc et al., in press).

	SUN	MON	TUE	WED	THU	FRI	SAT
	In-Proc./ Day 1	Training/ Day 2	Baseline/ Day 3	Testing 1/ Day 4	Testing 2/ Day 5	Testing 3/ Day 6	Recovery/ Day 7
00:00		Sleep	Sleep	Testing/**Dose**	Testing/**Dose**	Testing	Sleep
01:00				Simulator	Simulator	Simulator	
02:00				Testing	Testing	**Nap**	
03:00							
04:00				Testing/**Dose**	Testing/**Dose**	Wake/testing	
05:00				Simulator	Simulator	Simulator	
06:00		Wake/Shower	Wake/Shower	Shower	Shower	Shower	
07:00		Meal	Meal	Meal	Meal	Meal	Wake/showe r/Meal
08:00		Testing	Testing	Testing/**Dose**	Testing/**Dose**	Testing	Testing
09:00		Simulator	Simulator	Simulator	Simulator	Simulator	Simulator
10:00		Testing	Testing	Testing	Testing	Testing	Testing
11:00							
12:00	In-Process, EEG Hook-up, Lab Tour	Lunch	Lunch	Lunch	Lunch	Lunch	Med Exam
13:00		Simulator	Simulator	Simulator	Simulator	Simulator	Release
14:00		Testing	Testing	Testing	Testing	Testing	
15:00							
16:00		Break	Break	Break	Break	Break	
17:00		Simulator	Simulator	Simulator	Simulator	Simulator	
18:00		Testing	Testing	Testing	Testing	Testing	
19:00							
20:00	Dinner	Dinner	Dinner	Dinner	Dinner	Dinner	
21:00	PT	PT	PT	PT	PT	Sleep	
22:00	Sleep	Sleep	Testing	Testing	Testing		
23:00							

Appendix B

Aircrew coordination training performance evaluation checklist.

ACT Performance Evaluation Checklist

For use of this form, see ACT Aircrew Guide

CCO	BQ	Crew Coordination Objectives/Basic Qualities	Rating
1		Establish and Maintain Team Relationships	
	1	Establish and Maintain Team Leadership and Crew Climate	
2		Mission Planning and Rehearsal	
	2	Premission Planning and Rehearsal Accomplished	
	3	Application of Appropriate Decision Making Techniques	
3		Establish and Maintain Workload Levels	
	4	Prioritize Actions and Distribute Workload	
	5	Management of Unexpected Events	
4		Exchange Mission Information	
	6	Statements and Directives Clear, Timely, Relevant, Complete and Verified	
	7	Maintenance of Situational Awareness	
	8	Decisions and Actions Communicated and Acknowledged	
	9	Supporting Information and Actions Sought from Crew	
5		Cross-Monitor Performance	
	10	Crewmembers Actions Mutually Cross-Monitored	
	11	Supporting Information and Actions Offered by Crew	
	12	Advocacy and Assertion Practiced	
	13	Crew/Flight After-Action Reviews Accomplished	

Remarks: (Use continuation sheet(s) if necessary)

Notes:

Consult the ACT Aircrew Guide evaluation procedures and guidelines. Enter a summary rating (1-7) in the rating block for each ACT Crew Coordination Objective (CCO). Refer to the rating scale below.

Below Standards			Meets Standards			Exceeds Standards
1	2	3	4	5	6	7

<u>Appendix C</u>

<u>Performance bands (Caldwell et al., 1992, except DME).</u>

Variable (units)	Band limits					
	0%	20%	40%	60%	80%	100%
Heading (Degrees)	12.000-999.000	6.000- 12.000	3.000- 6.000	1.500- 3.000	0.750- 1.500	0.000- 0.750
Altitude (Feet)	140.000-999.000	70.000-140.000	35.000- 70.000	17.500- 35.000	8.750- 17.500	0.000- 8.750
Airspeed (Knots)	16.000-999.000	8.000- 16.000	4.000- 8.000	2.000- 4.000	1.000- 2.000	0.000- 1.000
Climb rate (Ft/min)	800.000-999.000	400.000-800.000	200.000-400.000	100.000-200.000	50.000-100.000	0.000-50.000
Pitch (Degrees)	6.000-999.000	3.000- 6.000	1.500- 3.000	0.750- 1.500	0.375- 0.750	0.000- 0.375
Roll (Degrees)	8.000-999.000	4.000- 8.000	2.000- 4.000	1.000- 2.000	0.500- 1.000	0.000- 0.500
Slip (Gs)	0.060-999.000	0.030- 0.060	0.015- 0.030	0.008- 0.015	0.004- 0.008	0.000- 0.004
Localizer (Dots)	3.800-999.000	1.900- 3.800	0.950- 1.950	0.475- 0.950	0.238- 0.475	0.000- 0.238
Glideslope (Dots)	3.800-999.000	1.900- 3.800	0.950- 1.950	0.475- 0.950	0.238- 0.475	0.000- 0.238
DME (Miles)	2.000- 999.000	1.000- 2.000	0.500- 1.000	0.250- 0.500	0.125- 0.250	0.000- 0.125

Appendix D

Descriptive statistics of performance scores by drug condition.

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Caffeine						
Altitude Score	24	56.09	35.91	92.00	64.3238	13.09411
Airspeed Score	24	38.00	60.00	98.00	80.4200	11.98840
DME arc Score	23	41.83	58.17	100.00	81.7213	12.55580
Total Maneuver Score	23	43.08	53.58	96.67	75.6562	9.62223
Altitude RMSE	24	72.83	8.58	81.41	34.5137	20.03778
Airspeed RMSE	24	6.60	.75	7.35	2.7621	1.90382
DME arc RMSE	23	.36	.02	.38	.1287	.09246
ACT Performance Evaluation Rating	16	4	1	5	2.88	1.204
Dextroamphetamine						
Altitude Score	24	47.64	44.19	91.83	70.1833	13.60446
Airspeed Score	24	90.70	9.30	100.00	82.4467	20.07777
DME arc Score	24	100.00	.00	100.00	84.1104	20.32620
Total Maneuver Score	24	73.75	19.81	93.57	78.9135	15.19882
Altitude RMSE	24	283.04	9.32	292.36	40.0992	57.57184
Airspeed RMSE	24	19.97	.29	20.26	3.3408	4.62418
DME arc RMSE	24	2.39	.02	2.41	.1729	.47803
ACT Performance Evaluation Rating	17	4	2	6	3.47	1.179
Modafinil						
Altitude Score	24	68.24	28.60	96.84	69.6313	16.14465
Airspeed Score	24	46.00	54.00	100.00	84.1267	13.00102
DME arc Score	24	49.82	50.18	100.00	90.7296	11.54123
Total Maneuver Score	24	29.81	69.14	98.95	81.4958	7.51683
Altitude RMSE	24	60.26	7.07	67.33	25.9350	13.64854
Airspeed RMSE	24	4.87	.27	5.14	2.0379	1.21484
DME arc RMSE	24	.27	.01	.28	.0708	.06164
ACT Performance Evaluation Rating	16	4	1	5	3.44	.964
Placebo						
Altitude Score	22	64.09	30.31	94.40	66.8000	16.69834
Airspeed Score	22	38.08	61.92	100.00	86.0223	11.68288
DME arc Score	22	36.36	63.64	100.00	85.5959	10.77892
Total Maneuver Score	22	31.33	61.76	93.09	79.4727	8.96755
Altitude RMSE	22	79.18	6.88	86.06	29.6591	17.34020
Airspeed RMSE	22	4.53	.25	4.78	1.9182	1.20203
DME arc RMSE	22	.28	.01	.29	.0968	.06890
ACT Performance Evaluation Rating	20	5	1	6	4.00	1.686